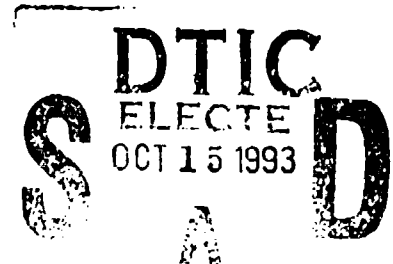




2

Principal Investigators Names: Robert C. Moore, Michael H. Cohen  
PI Institution: SRI International  
PI Phone Numbers: (415) 859-3448, (415) 859-5977  
PI E-mail Addresses: bmoore@ai.sri.com, mcohen@speech.sri.com  
Contract Titles: A Real-Time Spoken-Language System for Interactive Problem-Solving, Combining  
Linguistic and Statistical Technology for Improved Spoken Language Understanding  
Contract Numbers: N00014-90-C-0085, N00014-93-C-0142  
Reporting Period: 1 Oct 92 - 30 Sep 93

## Annual Report



### 1 Productivity Measures

Refereed papers submitted but not yet published: 1

Refereed papers published: 0

Unrefereed reports and articles: 6 (Most of these were conference papers for which 2- to 6-page extended abstracts were refereed.)

Books or parts thereof submitted but not yet published: 0

Books or parts thereof published: 0

Patents filed but not yet granted: 0

Patents granted: 0

Invited presentations: 1

Contributed presentations: 4

Honors received: 1 (Douglas Appelt was elected vice-president of the Association for Computational Linguistics for 1994.)

Prizes or awards received: 0

Promotions obtained: 1 (Patti Price was promoted to Director of SRI's Speech Research Program.)

Graduate students supported  $\geq$  25% of full time: 3

Post-docs supported  $\geq$  25% of full time: 1

Minorities supported: 0

This document has been approved  
for public release and sale; its  
distribution is unlimited

93-24052



Principal Investigators Names: Robert C. Moore, Michael H. Cohen  
 PI Institution: SRI International  
 PI Phone Numbers: (415) 859-3448, (415) 859-5977  
 PI E-mail Addresses: bmoore@ai.sri.com, mcohen@speech.sri.com  
 Contract Titles: A Real-Time Spoken-Language System for Interactive Problem-Solving, Combining  
 Linguistic and Statistical Technology for Improved Spoken Language Understanding  
 Contract Numbers: N00014-90-C-0085, N00014-93-C-0142  
 Reporting Period: 1 Oct 92 - 30 Sep 93

## 2 Summary of Technical Progress

Under this effort, SRI has developed spoken-language technology for interactive problem solving, featuring real-time performance for up to several thousand word vocabularies, high semantic accuracy, habitability within the domain, and robustness to many sources of variability. Although the technology is suitable for many applications, efforts to date have focussed on developing an Air Travel Information System (ATIS) prototype application. SRI's ATIS system has been evaluated in four ARPA benchmark evaluations, and has consistently been at or near the top in performance. These achievements are the result of SRI's technical progress in speech recognition, natural-language processing, and speech and natural-language integration.

SRI's DECIPHER is a continuous-speech, speaker-independent speech recognizer based on tied-mixture, gender-dependent hidden Markov model (HMM) technology. It uses six cepstra- and energy-based features generated from a filterbank computed via fast Fourier transforms and high-pass filtering in the log-spectral-energy domain. Noise-robust modeling and spectral normalization algorithms are used to improve robustness to channel variation. Pronunciation variability is modeled through probabilistically pruned linguistic rules. Cross-word acoustic and phonological models are used to model coarticulation. Recognizers trained separately on male and female speech are run in parallel, and backed-off bigram and trigram language models are used to reduce perplexity. Current research focuses on improving the consistency of recognition hypotheses by reducing frame-to-frame independence assumptions, adapting acoustic models to speakers, and improving the estimation and use of statistical language models.

SRI's research on natural-language understanding for spoken-language systems has proceeded along two lines. The shorter-term line of research has focused on the Template Matcher, a module that constructs database queries by searching the user input for key words and phrases characteristic of the most common query types for a given task, ignoring parts of the input that it does not understand. This approach is robust to unanticipated variation in phrasing, but is limited in its ability to extract information that depends on structural relationships among phrases. A longer-term effort is focused on more sophisticated syntactic and semantic analysis of the input, using a unification-grammar-based natural-language processing system called GEM-

✓	
□	
□	
PR A266046	
ACQUISITION CENTER	
Dist	Acad and/or Special
A-1	

INI. GEMINI is capable of analyzing more complex semantic relationships than the Template Matcher, but is more fragile, by itself, to unanticipated variation in query phrasing. To get the benefits of both approaches, the ATIS system we currently use in benchmark evaluations incorporates both GEMINI and the Template Matcher, by first attempting to construct a complete analysis of a query using GEMINI, and falling back on the Template Matcher if that fails. That way, GEMINI gets a chance to give an exact analysis of the input before the Template Matcher attempts an approximate one.

The integration of speech and language processing in SRI's current ATIS system is a serial connection, using the robustness of the Template Matcher to accommodate recognition errors. This resulted in a system where only 21.6% of utterances in the November 1992 ATIS SLS benchmark test failed to receive a correct answer, even though there was at least one recognition error on 33.8% of the utterances. Longer-term research is exploring the use of the GEMINI system to guide the recognizer to favor more semantically meaningful recognition hypotheses. Our current research in this area focuses on maintaining robustness by making use of information provided by the natural-language system even when it fails to obtain a complete semantic analysis.

Another area of research in speech and natural-language integration in which we have achieved significant results is the detection and correction of verbal repairs—cases of self-editing where the speaker intends part of what was said to be overridden by subsequent speech; for example, *Show me flights (that lea-) that arrive before seven pm.* In this effort, we have developed a notation for describing and annotating repairs, analyzed the repairs occurring in a 10,000-utterance training set of ATIS data, and developed preliminary methods to recognize and correct repairs by combining string matching with acoustic and natural-language information sources. In addition, we have incorporated a component based on those methods into the GEMINI system.

During current reporting period of the project we have:

- Participated in ARPA-sponsored ATIS benchmarks: SRI achieved a 9.1% word-error rate for speech recognition, a 23.6% weighted utterance error for natural language understanding, and 33.2% weighted utterance error for speech understanding.
- Prepared and delivered to ARPA demonstrations of spoken-language technology (voice banking, ATIS, and dictation) along with associated online background material.
- Carried out experiments suggesting a potential 30–40% reduction in word-error rate by incorporating natural-language constraints from GEMINI into DECIPHER.
- Increased the speed of the GEMINI parser by a factor of four by improved handling of semantic selectional restrictions.

- Developed a method for constructing telephone acoustic models from a high quality speech corpus.
- Developed a new method for speaker adaptation based on Bayesian reestimation of tied-gaussian mixture weights.
- Developed trigram-based language models for the ATIS task.
- Revised the method for generating progressive search lattices to facilitate the use of natural-language knowledge sources to constrain speech recognition.
- Developed techniques for using the Template Matcher to constrain speech recognition.
- Collected spontaneous speech data in the ATIS domain from 47 speakers for 190 scenarios.

Principal Investigators Names: Robert C. Moore, Michael H. Cohen  
PI Institution: SRI International  
PI Phone Numbers: (415) 859-3448, (415) 859-5977  
PI E-mail Addresses: bmoore@ai.sri.com, mcohen@speech.sri.com  
Contract Titles: A Real-Time Spoken-Language System for Interactive Problem-Solving, Combining  
Linguistic and Statistical Technology for Improved Spoken Language Understanding  
Contract Numbers: N00014-90-C-0085, N00014-93-C-0142  
Reporting Period: 1 Oct 92 - 30 Sep 93

### 3 Publications, Presentations, and Reports

Bear, J., J. Dowding, E. Shriberg, and P. Price, "A System for Labeling Self-Repairs in Speech." Technical Note 522, SRI International, Menlo Park, California (February 1993).

Dowding, J., J. M. Gawron, D. Appelt, J. Bear, L. Cherny, R. Moore, and D. Moran, "Gemini: A Natural Language System for Spoken-Language Understanding," paper presented at the ARPA Workshop on Human Language Technology, Plainsboro, New Jersey, March 1993.

Dowding, J., J. M. Gawron, D. Appelt, J. Bear, L. Cherny, R. Moore, and D. Moran, "Gemini: A Natural Language System for Spoken-Language Understanding," paper presented at the 31st Annual Meeting of the Association for Computational Linguistics, Columbus, Ohio, June 1993 (*Proceedings*, pp. 54-61).

Jackson, E., "Integrating Two Complementary Approaches to Spoken Language Understanding," paper presented at ICSLP 92, 1992 International Conference on Spoken Language Processing, Banff, Alberta, Canada, October 1992 (*Proceedings*, pp. 333-336).

Moore, R., "Integration of Speech with Natural-Language Understanding," invited presentation at the National Academy of Sciences Colloquium on Human-Machine Communication by Voice, The Arnold and Mabel Beckman Center, Irvine, California, February 1993.

Neumeyer, L., V. Digalakis, and M. Weintraub, "Training Issues and Channel Equalization Techniques for the Construction of Telephone Acoustic Models Using a High-Quality Speech Corpus," to appear in *IEEE Transaction on Speech and Audio*.

Price, P., and R. Moore, "A Real-Time Spoken-Language System for Interactive Problem Solving," Final Report, Contract N00014-90C-0085, SRI Project 8900, SRI International, Menlo Park, California (May 1993).

Wade, E., E. Shriberg, and P. Price, "User Behaviors Affecting Speech Recognition," paper presented at ICSLP 92, 1992 International Conference on Spoken Language Processing, Banff, Alberta, Canada, October 1992 (*Proceedings*, pp. 995-998).

Principal Investigators Names: Robert C. Moore, Michael H. Cohen  
PI Institution: SRI International  
PI Phone Numbers: (415) 859-3446, (415) 859-5977  
PI E-mail Addresses: bmoore@ai.sri.com, mcohen@speech.sri.com  
Contract Titles: A Real-Time Spoken-Language System for Interactive Problem-Solving, Combining  
Linguistic and Statistical Technology for Improved Spoken Language Understanding  
Contract Numbers: N00014-90-C-0085, N00014-93-C-0142  
Reporting Period: 1 Oct 92 - 30 Sep 93

## **4 Transitions and DoD Interactions**

SRI delivered to ARPA demonstration versions of three prototype systems (voice banking, ATIS, and dictation) for promotion and for the education of potential applications developers, users, and funders. SRI's ATIS system has been delivered to NIST for collection of training and test data for benchmark evaluations, and we have trained NIST personnel in data collection. Finally, SRI participated in the ARPA-sponsored Spoken Language Technology Applications Day in Washington, DC, in April 1993, demonstrating applications of our spoken-language technology.

Principal Investigators Names: Robert C. Moore, Michael H. Cohen  
PI Institution: SRI International  
PI Phone Numbers: (415) 859-3448, (415) 859-5977  
PI E-mail Addresses: bmoore@sri.com, mcohen@speech.sri.com  
Contract Titles: A Real-Time Spoken-Language System for Interactive Problem-Solving, Combining  
Linguistic and Statistical Technology for Improved Spoken Language Understanding  
Contract Numbers: N00014-90-C-0085, N00014-93-C-0142  
Reporting Period: 1 Oct 92 - 30 Sep 93

## 5 Software and Hardware Prototypes

We have developed an Air Travel Information System (ATIS) as a prototype application of spoken-language understanding. This system enables a user to retrieve airline schedules, fares, and related information by means of spoken natural-language queries. We have evaluated this system in four ATIS benchmark evaluations, and we have also used it for data collection.

We have also developed GEMINI, a parsing and semantic interpretation system based on unification grammar. GEMINI first applies syntactic, semantic, and lexical rules bottom-up, using an all-paths "constituent" parser to populate a chart with edges containing syntactic and semantic information. Then, a second "utterance" parser is used to apply a second set of syntactic and semantic rules that are required to span the entire utterance. If no semantically acceptable utterance-spanning edges are found during this phase, a component to recognize and correct verbal repairs is applied. When an acceptable interpretation is found, a set of parse preferences is used to choose a single best interpretation from the chart to be used for subsequent processing. Quantifier scoping rules are applied to this best interpretation to produce a scoped logical form.

In our efforts to commercialize the spoken-language systems technology developed in part under this effort, we have developed a Trader's Workstation for a large Wall Street firm, which they are currently evaluating.